EFFICIENT ESTIMATION OF SIMULTANEOUS EQUATIONS WITH AUTO-REGRESSIVE ERRORS BY INSTRUMENTAL VARIABLES

Ray C. Fair *

I Introduction

THE purpose of this paper is to point out how the efficient instrumental-variables technique discussed by Brundy and Jorgenson (1971) can be modified to take into account auto-regressive properties of the error terms. The limited-information and full-information estimators proposed in this paper are consistent and have the same asymptotic distributions as the limited-information and full-information maximum likelihood estimators, respectively.

The full-information estimation of simultaneous equations models with auto-regressive errors has been discussed by Sargan (1961), Hendry (1971), Chow and Fair (1973), and Dhrymes (1971). Sargan originally proposed the full-information maximum likelihood estimation of such models, and Hendry and Chow and Fair have recently developed computationally feasible methods for obtaining the maximum likelihood estimates. Hendry considered only the case of completely unrestricted auto-regressive coefficient matrices (i.e., no zero elements), whereas Chow and Fair considered the case of restricted auto-regressive coefficient matrices as well. Dhrymes has recently proposed the three-stage least squares estimator of simultaneous equations models with auto-regressive errors. Dhrymes also considered only the case of completely unrestricted auto-regressive coefficient matrices.

The limited-information estimation of simultaneous equations models with auto-regressive errors has been discussed by Sargan (1961), Amemiya (1966), and Fair (1970), among others. Sargan proposed the limited-information maximum likelihood estimation of such models, and Amemiya and Fair considered various two-stage least squares estimators of such models. Most of the work on limited-information estimators has been concerned with the case of diagonal auto-regressive coefficient matrices.

Brundy and Jorgenson's criticism of the two- and three-stage least squares estimators, namely, that the first stage involves estimating reduced form equations with a very large number of variables included in them, holds even more so for models with auto-regressive errors. For these models, the reduced form equations include not only all of the predetermined variables in the system but also all of the lagged endogenous and lagged predetermined variables. In fact, one of the main purposes of the work by Fair (1970) was to suggest ways in which the number of variables used in the first stage regressions of two-stage least squares might be decreased with perhaps small loss of asymptotic efficiency. The advantage of the instrumental-variables techniques proposed in the Brundy-Jorgenson paper and in this paper is that the first stage regressions need not be run.

II The Model

The model to be estimated is [1]

$$Y\Gamma + XB = U, \qquad (1)$$

where Y is a $n \times p$ matrix of endogenous variables, X is a $n \times q$ matrix of predetermined variables, U is a $n \times p$ matrix of error terms, and Γ and B are $p \times p$ and $q \times p$ coefficient matrices, respectively. The X matrix may include lagged endogenous variables as well as exogenous variables. n is the number of observations. As distinct from the Brundy-Jorgenson paper, it is assumed here that the error terms in U follow a m^{th} order auto-regressive process:

$$U = U_{-1}R^{(1)} + \ldots + U_{-m}R^{(m)} + E, \qquad (2)$$

Received for publication October 1, 1971. Revision accepted for publication June 13, 1972.

* I thank Gregory C. Chow and Phoebus J. Dhrymes for helpful discussions regarding certain parts of this paper.

[1] The notation here follows as closely as possible the notation in Brundy and Jorgenson (1971).

where the $R^{(k)}$ matrices are $p \times p$ coefficient matrices, E is a $n \times p$ matrix of error terms, and the subscripts denote lagged values of the terms of U. Combining (1) and (2) yields

$$Y\Gamma + XB = Y_{-1}\Gamma R^{(1)} + X_{-1}BR^{(1)} + \ldots + Y_{-m}\Gamma R^{(m)} + X_{-m}BR^{(m)} + E. \tag{3}$$

From (3) the reduced form for Y is

$$Y = -XB\Gamma^{-1} + Y_{-1}\Gamma R^{(1)}\Gamma^{-1} + X_{-1}BR^{(1)}\Gamma^{-1} + \ldots + Y_{-m}\Gamma R^{(m)}\Gamma^{-1} + X_{-m}BR^{(m)}\Gamma^{-1} + E\Gamma^{-1}, \tag{4}$$

or

$$Y = Q\Pi + V, \tag{5}$$

where $V = E\Gamma^{-1}$, $Q = [X\ Y_{-1}\ X_{-1} \ldots Y_{-m}\ X_{-m}]$, and Π is partitioned according to Q.

It is convenient to write the structural equations in (1) in the form:

$$y_j = Z_j\delta_j + u_j, \quad j = 1, 2, \ldots, p, \tag{6}$$

where

$$Z_j = [Y_j X_j], \quad \delta_j = \begin{bmatrix} \gamma_j \\ \beta_j \end{bmatrix}.$$

As in Brundy and Jorgenson (1971, p. 208), y_j is a vector of observations on the j^{th} column of Y whose structural coefficient has been normalized to one, Y_j is a matrix of observations on the other endogenous variables included in the equation, X_j is a matrix of observations on the predetermined variables included directly in the equation, u_j is the j^{th} column of U, and γ_j and β_j are structural coefficient vectors corresponding to Y_j and X_j respectively. The p equations in (6) can be combined to yield

$$y = Z\delta + u, \tag{7}$$

where

$$y = \begin{bmatrix} y_1 \\ y_2 \\ \cdot \\ \cdot \\ \cdot \\ y_p \end{bmatrix}, \ Z = \begin{bmatrix} Z_1 & 0 & \ldots & 0 \\ 0 & Z_2 & \ldots & 0 \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ 0 & 0 & \ldots & Z_p \end{bmatrix}, \ \delta = \begin{bmatrix} \delta_1 \\ \delta_2 \\ \cdot \\ \cdot \\ \cdot \\ \delta_p \end{bmatrix}, \ u = \begin{bmatrix} u_1 \\ u_2 \\ \cdot \\ \cdot \\ \cdot \\ u_p \end{bmatrix}.$$

In order to implement the instrumental variables estimator in the auto-regressive case, it is necessary to transform (7) so that the error term on the right-hand side is e rather than u, where

$$e = \begin{bmatrix} e_1 \\ e_2 \\ \cdot \\ \cdot \\ \cdot \\ e_p \end{bmatrix},$$

e_j being the j^{th} column of E. This transformation is:

$$y - (R^{(1)\prime} \mathrm{x} I)y_{-1} - \ldots - (R^{(m)\prime} \mathrm{x} I)y_{-m} = [Z - (R^{(1)\prime} \mathrm{x} I)Z_{-1} \ldots - (R^{(m)\prime} \mathrm{x} I)Z_{-m}]\delta + e, \tag{8}$$

or

$$\bar{y} = \bar{Z}\delta + e, \tag{9}$$

where I is an $n \times n$ identity matrix and the subscripts on y and Z denote lagged values.

III The Full-information Estimator

The basic idea of the Brundy-Jorgenson paper is that if a set of instrumental variables can be found that is based on a consistent estimate of Π, then using this set of instrumental variables will result in asymptotically efficient estimates (within the class of either limited-information or full-information methods). In the present case, Π in (5) is a function of the $R^{(k)}$ matrices as well as of Γ and B. Consequently, if consistent estimates of Γ, B, and the $R^{(k)}$ matrices are available, then a consistent estimate of Π in (5) is available. The equations in (5) can then be used to generate consistent predictions of the endogenous variables. Consistent estimates of Γ, B, and the $R^{(k)}$ matrices can also be used to obtain a consistent estimate of the variance-covariance matrix, Σ, of the error terms E in (3).

Assume, therefore, that initial consistent estimates of Γ, B, and the $R^{(k)}$ matrices are available [2] so that a consistent estimate of Π

[2] It will be seen in section VI how initial consistent estimates of these matrices can be obtained.

in (5) is available. The matrix Z consists of current endogenous variables as well as of predetermined variables. Since a consistent estimate of Π is assumed to be available, (5) can be used to generate consistent predictions of the endogenous variables in the model. Let Z^* denote the matrix Z except for the replacement of the current endogenous variables in Z by their predicted values from (5). Let $\bar{y}^*$ and $\bar{Z}^*$ denote the matrices $\bar{y}$ and $\bar{Z}$, respectively, except for the use of consistent estimates of the $R^{(k)}$ matrices rather than the actual matrices to transform the variables. Also, let $\bar{Z}^{**}$ denote the matrix $\bar{Z}^*$ except for the replacement of Z by Z^*, and let $W = (\Sigma^{*-1} \mathrm{x} I)\, \bar{Z}^{**}$, where Σ^* is a consistent estimate of Σ. Then the "full-information instrumental variables efficient"[3] estimator in the auto-regressive case (say, *FIVER*) can be defined to be:

$$d = (W'\bar{Z}^*)^{-1}W'\bar{y}^*. \tag{10}$$

It is easy to show that the *FIVER* estimator is consistent. From (8) or (9) and the definition of $\bar{y}^*$ and $\bar{Z}^*$, it follows that

$$\bar{y}^* = \bar{Z}^*\delta + ((R^{(1)\prime} - R^{*(1)\prime})\mathrm{x}I)u_{-1} + \ldots + ((R^{(m)\prime} - R^{*(m)\prime})\mathrm{x}I)u_{-m} + e, \tag{11}$$

where the $R^{*(k)}$ matrices are consistent estimates of the $R^{(k)}$ matrices. Substituting (11) into (10) yields:

$$d = \delta + (W'\bar{Z}^*)^{-1}W'[((R^{(1)\prime} - R^{*(1)\prime})\mathrm{x}I)u_{-1} + \ldots + ((R^{(m)\prime} - R^{*(m)\prime})\mathrm{x}I)u_{-m}] + (W'\bar{Z}^*)^{-1}W'e. \tag{12}$$

Assuming that plim $n\ (W'\bar{Z}^*)^{-1}$ exists and is finite, it follows that plim $d = \delta$, since plim $n^{-1} W'e = 0$ because of the inclusion in W of only predetermined variables or linear functions of predetermined variables and since the $R^{(k)}$ matrices are consistently estimated.[4]

It is also easy to see that the *FIVER* estimator is asymptotically efficient if the $R^{(k)}$ matrices are known with certainty. In this case the model in (9) is merely a standard simultaneous equations model in $\bar{y}$ and $\bar{Z}$. Brundy and Jorgenson have shown that the full-information instrumental-variables estimator based on instruments generated from a consistent estimate of the reduced form matrix Π and the three-stage least squares estimator have the same asymptotic distribution. Now, the formula for the *FIVER* estimator in (10) is the instrumental-variables analogue to the three-stage least squares formula (30) in Dhrymes (1971).[5] Therefore, since the three-stage least squares estimator of the standard simultaneous equations model is asymptotically efficient, the *FIVER* estimator is asymptotically efficient in the case of known $R^{(k)}$ matrices.

For the case in which the $R^{(k)}$ matrices are unknown and must be estimated, Dhrymes (1971) has shown for the three-stage least squares estimator that if one iterates back and forth between estimates of δ and estimates of the $R^{(k)}$ matrices and if convergence is reached, then asymptotically, the set of equations that is solved by this procedure is the same set of equations that the full-information maximum likelihood estimator satisfies. This conclusion also holds for the *FIVER* estimator, because asymptotically formula (10) for the *FIVER* estimator is the same as Dhrymes' formula (30) for the three-stage least squares estimator. Therefore, the *FIVER* estimator, based on iterating back and forth between estimates of δ and estimates of the $R^{(k)}$ matrices[6] until convergence is reached, has the same asymptotic distribution as the full-information maximum likelihood estimator.

[3] Brundy and Jorgenson (1971), p. 214.

[4] See Dhrymes (1971), Lemma 8, for a detailed proof of the consistency of the three-stage least squares estimator of the first-order auto-regressive model. A detailed proof in the present case would proceed in a similar manner as Dhrymes' proof.

[5] Dhrymes actually considered only the case of a first-order auto-regressive process, but it is easy to generalize his arguments and formulas to higher-order processes. Likewise, although Dhrymes considered only the case of completely unrestricted auto-regressive coefficient matrices, his formula (30) is valid for the case of restricted matrices as well.

In the notation of this paper, Dhrymes' formula (30) generalized to higher-order auto-regressive processes is

$$d = (W'\bar{Z}^{**})^{-1} W'\bar{y}^*,$$

where $\bar{Z}^{**}$ (which is also included in the definition of W) for the three-stage least squares case differs from the $\bar{Z}^{**}$ for the instrumental-variables case in that the predictions of the endogenous variables, which are a part of $\bar{Z}^{**}$, are based on first-stage, reduced-form regressions rather than on generated predictions from (5) using any consistent estimate of π. It can be seen, using the fact that both the three-stage least squares estimator and the instrumental-variables estimator are based on consistent estimates of π and the fact that the predicted values of the reduced-form error terms for the three-stage least squares estimator are orthogonal to all of the predetermined variables in the model, that formula (10) and Dhrymes' formula (30) are asymptotically the same.

[6] The estimation of the $R^{(k)}$ matrices is discussed in section V.

Ignoring the stochastic nature of the estimates of the $R^{(k)}$ matrices, the asymptotic variance-covariance matrix of the *FIVER* estimator is:

$$\text{asy}\cdot\text{var-cov } d = n^{-1} \text{ plim } n \; (W'\bar{Z}^*)^{-1} W' ee'W \; (\bar{Z}^{*\prime}W)^{-1}. \tag{13}$$

From the fact that Σ^* is a consistent estimate of Σ, that $\bar{Z}^{**}$ differs from $\bar{Z}^*$ merely by the replacement of the endogenous variables in $\bar{Z}^*$ by predictions of the endogenous variables based on a consistent estimate of Π in (5), that the variance-covariance matrix of e is $\Sigma x I$, and that $W = (\Sigma^{*-1} x I)\bar{Z}^{**}$, (13) reduces to n^{-1} plim n $(W'\bar{Z}^*)^{-1}$. The asymptotic variance-covariance matrix of d can thus be estimated as $(W'\bar{Z}^*)^{-1}$, although this estimate ignores the stochastic nature of the estimates of the $R^{(k)}$ matrices.

IV The Limited-information Estimator

In this section the limited-information case will be analyzed under the assumption that the $R^{(k)}$ matrices are diagonal. A brief description of how one can estimate models with non-diagonal $R^{(k)}$ matrices by limited-information techniques is presented in section VII.

If the $R^{(k)}$ matrices are diagonal, then equation (6) can be transformed as:

$$y_j - r_{jj}^{(1)} y_{j-1} - \ldots - r_{jj}^{(m)} y_{j-m} = (Z_j - r_{jj}^{(1)} Z_{j-1} - \ldots - r_{jj}^{(m)} Z_{j-m})\delta_j + e_j, \; j = 1, 2, \ldots, p, \tag{14}$$

or

$$\bar{y}_j = \bar{Z}_j \delta_j + e_j, \tag{15}$$

where the subscripts on y_j and Z_j denote lagged values and where $r_{jj}^{(k)}$ is the j^{th} diagonal element of $R^{(k)}$ $(k = 1, \ldots, m)$. Let Z^*_j denote the matrix Z_j except for the replacement of the current endogenous variables in Z_j by their predicted values from (5). Let $\bar{y}^*_j$ and $\bar{Z}^*_j$ denote the matrices $\bar{y}_j$ and $\bar{Z}_j$, respectively, except for the use of consistent estimates of the $r_{jj}^{(k)}$ coefficients rather than the actual coefficients to transform the variables. Also, let W_j denote the matrix $\bar{Z}^*_j$ except for the replacement of Z_j *by* Z^*_j. Then the "limited-information instrumental variables efficient"[7] estimator in the auto-regressive case (say, *LIVER*) is:

$$d_j = (W'_j\bar{Z}^*_j)^{-1} W'_j \bar{y}^*_j. \tag{16}$$

The discussion of the asymptotic properties of the *LIVER* estimator is similar to the discussion of the asymptotic properties of the *FIVER* estimator and need only be briefly elaborated on here. The *LIVER* estimator is consistent, and within the class of limited-information estimators, the estimator is asymptotically efficient if the $r_{jj}^{(k)}$ coefficients are known with certainty. For the case in which the $r_{jj}^{(k)}$ coefficients are unknown and must be estimated, Amemiya (1966) in equations (19) and (20) has presented the two-stage least squares analogue of the limited-information maximum likelihood estimator and has shown that the two estimators have the same asymptotic distribution. Equations (19) and (20) in Amemiya (1966) can be solved by iterating back and forth between estimates of the structural coefficients and estimates of the auto-regressive coefficient.[8] Asymptotically, formula (16) for the *LIVER* estimator is the same as Amemiya's formula (19) for the two-stage least squares estimator.[9] Therefore, the *LIVER* estimator, based on iterating back and forth between estimates of δ_j and estimates of the $r_{jj}^{(k)}$ coefficients [10] until convergence is reached, has the same asymptotic distribution as the limited-information maximum likelihood estimator. Ignoring the stochastic nature of the estimates of the $r_{jj}^{(k)}$ coefficients, the asymptotic variance-covariance matrix of the *LIVER* estimator is $n^{-1}\sigma_{jj}$ plim $n(W'_j\bar{Z}^*_j)^{-1}$, which can be estimated as $\sigma^*_{jj}(W'_j\bar{Z}^*_j)^{-1}$, where σ^*_{jj} is the j^{th} diagonal element of Σ^*.

V Estimates of the $R^{(k)}$ Matrices

Given consistent estimates of the Γ and B matrices, consistent estimates of the error

[7] Brundy and Jorgenson (1971), p. 211.

[8] The scanning and iterative procedures discussed in Fair (1970) are two ways of solving equations (19) and (20) in Amemiya (1966), although the basic estimator discussed in Fair (1970) is equivalent to Amemiya's two-stage least squares analogue of the limited-information maximum likelihood estimator only if all of the predetermined, lagged predetermined, and lagged endogenous variables are used as regressors in the first stage regressions. Otherwise, efficiency is lost if not all of these variables are used as regressors in the first stage regressions.

[9] Amemiya considered the case of a first-order auto-regressive process, but it is easy to generalize his formulas to higher order processes.

[10] The estimation of the $r_{jj}^{(k)}$ coefficients is discussed in section V.

matrices $U, U_{-1}, \ldots, U_{-m}$ can be obtained from the current and lagged versions of (1). Let U^* denote any consistent estimate of U, and let $\bar{U}^*$ denote any consistent estimate of $\bar{U}$, where $\bar{U} = (U_{-1} \ldots U_{-m})$. Also, let $\bar{R}' = (R^{(1)\prime} \ldots R^{(m)\prime})$ and write (2) as

$$U = \bar{U}\bar{R} + E. \tag{17}$$

Now, for known values of U and $\bar{U}$, (17) can be interpreted as a Zellner "seemingly unrelated regression" model unless the $\bar{R}$ matrix is completely unrestricted, in which case (17) is merely the standard multivariate linear regression model. Since consistent estimates of U and $\bar{U}$ are available, for the full-information case $\bar{R}$ can be estimated as

$$\bar{R}^* = (\bar{U}^{*\prime}\Sigma^{-1}\bar{U}^*)^{-1}\bar{U}^{*\prime}\Sigma^{*1-}U^*, \tag{18}$$

where Σ^* is a consistent estimate of Σ. For the case in which $\bar{R}$ is completely unrestricted, the full-information estimator is merely $(\bar{U}^{*\prime}\bar{U}^*)^{-1}\bar{U}^{*\prime}U^*$. This later case is the case analyzed by Hendry (1971) and Dhrymes (1971). For the limited-information case, the (diagonal) elements of the $R^{(k)}$ matrices can be estimated by merely regressing each column of U^* on the corresponding columns of $U^*_{-1}, \ldots, U^*_{-m}$. For the limited-information case, information about Σ^* is ignored.

If one iterates back and forth between estimates of the structural coefficients and estimates of the $R^{(k)}$ matrices and if convergence is reached, then, as discussed above, the limited-information and full-information estimators of the $R^{(k)}$ matrices have the same asymptotic distributions as the limited-information and full-information maximum likelihood estimators respectively. Convergence is, of course, not guaranteed from iterating. It is interesting to note that for the full-information maximum likelihood case, iterating back and forth between the estimates of the structural coefficients and the estimates of the $R^{(k)}$ matrices is guaranteed to converge provided that the separate maximization problems can be solved — see Chow and Fair (1973).

Ignoring the stochastic nature of the estimates of the structural coefficients, the asymptotic variance-covariance matrix of $\bar{R}^*$ can be estimated as $(\bar{U}^{*\prime}\Sigma^{*-1}\bar{U}^*)^{-1}$, or as $(\bar{U}^{*\prime}\bar{U}^*)^{-1}$ if $\bar{R}$ is completely unrestricted. In the limited information case, an estimate of the variance-covariance matrix of the $r_{jj}^{(k)}$ coefficient estimates can be taken to be the estimate of the variance-covariance matrix computed from each of the least squares regressions.

VI Obtaining Initial Consistent Estimates

There are many ways in which initial consistent estimates of Γ, B, and the $R^{(k)}$ matrices can be obtained. One general technique is as follows: Treat all lagged endogenous variables (as well as endogenous variables) as endogenous, and estimate each equation of (1) by instrumental variables ignoring the auto-regressive properties of the error terms. This will result in consistent estimates of Γ and B as long as only exogenous and lagged exogenous variables are used as instruments. Use these consistent estimates to compute consistent estimates of the residuals $U, U_{-1}, \ldots, U_{-m}$. Then for each equation, regress the unlagged estimated residuals on the appropriate lagged estimated residuals. The set of lagged estimated residuals will, in general, include both lagged estimated residuals of the particular equation being estimated as well as lagged estimated residuals of other equations of the model. This procedure will yield consistent estimates of the $R^{(k)}$ matrices since the residuals are consistently estimated. In special cases (such as diagonal $R^{(k)}$ matrices) there are, of course, other techniques that can be used to obtain initial consistent estimates. For example, in the first-order case with a diagonal $R^{(1)}$ matrix, the technique described in Fair (1970) can be used.[11]

[11] Dhrymes, Berner, and Cummins (1970) have also considered the estimation of the first-order auto-regressive model with a diagonal $R^{(1)}$ matrix. The estimator that they propose is similar to, but is not, a *LIVER* estimator. Dhrymes, Berner, and Cummins first obtain consistent estimates of Γ and B in (1) by an instrumental-variables technique treating lagged endogenous variables as endogenous. They then use these estimates in the reduced form of (1) — ignoring the auto-regressive process of U — and obtain a set of instrumental variables by dynamic simulation (i.e., using generated values of the lagged endogenous variables as opposed to the actual values). They also use the estimates of Γ and B to estimate U and U_{-1} and from the estimates of U and U_{-1} to obtain estimates of the diagonal elements of $R^{(1)}$ by ordinary least squares. They then use the set of instrumental variables and the estimates of the elements of $R^{(1)}$ to obtain new estimates of Γ and B. Johnston has shown that the estimator is not asymptotically efficient within the class of limited-information estimators.

VII Limited-information Estimation of Models with Nondiagonal $R^{(k)}$ Matrices

In this section it will be shown how limited information techniques can be used to estimate models with nondiagonal $R^{(k)}$ matrices. Assume without the loss of generality that the following equation is to be estimated:

$$y_j = Z_j\delta_j + u_j, \tag{19}$$

where

$$Z_j = [Y_j X_j],\ \delta_j = \begin{bmatrix} \gamma_j \\ \beta_j \end{bmatrix},\quad u_j = r_{jj}{}^{(1)}u_{j-1} + r_{ij}{}^{(1)}u_{i-1} + e_j,$$

$r_{jj}{}^{(1)}$ and $r_{ij}{}^{(1)}$ being elements of $R^{(1)}$. Equation (19) can be rewritten as

$$\begin{aligned} y_j - r_{ij}{}^{(1)}u_{i-1} &= Z_j\delta_j + r_{jj}{}^{(1)}\, u_{j-1} + e_j \\ &= Z_j\delta_j + r_{jj}{}^{(1)}y_{j-1} - r_{jj}{}^{(1)}Z_{j-1}\delta_j + e_j \end{aligned} \tag{20}$$

or

$$\bar{y}_j = \bar{Z}_j\delta_j + e_j, \tag{21}$$

where

$$\bar{y}_j = y_j - r_{ij}{}^{(1)}u_{i-1} - r_{jj}{}^{(1)}y_{j-1},$$
$$\bar{Z}_j = Z_j - r_{jj}{}^{(1)}Z_{j-1}.$$

Equation (21) is in a form like (15) except for the inclusion of the $-\,r_{ij}{}^{(1)}u_{i-1}$ term in $\bar{y}_j$. If consistent estimates of r_{ij} and the residual vector u_{i-1} are available, however, then the estimation of (21) by the *LIVER* technique can proceed like the estimation of (15). All that has been done is the subtraction of a consistent estimate of $r_{ij}{}^{(1)}u_{i-1}$ from $\bar{y}_j$ in (15).

VIII Conclusion

One of the main advantages of the estimators proposed in this paper is that first-stage, reduced-form regressions do not have to be run. For the single-equation case, a disadvantage is that an entire model must be specified and consistently estimated in order to obtain efficient estimates of any single equation. In at least some practical applications this may be a serious disadvantage, and for these cases one might wish to resort to a less efficient estimator like the two-stage least squares estimator proposed in Fair (1970), which does not require the specification and estimation of the entire model and does not require that all of the predetermined, lagged predetermined, and lagged endogenous variables be used regressors in the first stage regressions.

Since in the full-information case it is now feasible to estimate models with auto-regressive errors by the maximum likelihood method, it might be desirable to attempt to estimate a model by full-information maximum likelihood (*FIML*) before resorting to the *FIVER* estimator. As discussed in Chow (1964), there are some methodological reasons for preferring the *FIML* estimator over other asymptotically efficient estimators. It is possible, however, that the *FIVER* estimator will be able to handle larger models than will the *FIML* estimator.

REFERENCES

Amemiya, T .,"Specification Analysis in the Estimation of Parameters of a Simultaneous Equation Model with Autoregressive Residuals," *Econometrica*, XXXIV (Apr. 1966), 283–306

Brundy, J. M., and D. W. Jorgenson, "Efficient Estimation of Simultaneous Equations by Instrumental Variables," this REVIEW, LIII (Aug. 1971), 207–224.

Chow, G. C., "A Comparison of Alternative Estimators for Simultaneous Equations," *Econometrica*, XXXII (Oct. 1964), 532–553.

Chow, G. C., and R. C. Fair, "Maximum Likelihood Estimation of Linear Equation Systems with Auto-Regressive Residuals," *Annals of Economic and Social Measurement* (forthcoming, Jan. 1973).

Dhrymes, P. J., "Full Information Estimation of Dynamic Simultaneous Equations Models with Autoregressive Errors," Discussion Paper no. 203, The Wharton School of Finance and Commerce, University of Pennsylvania (March, 1971).

Dhrymes, P. J., R. Berner, and R. Cummins, "Limited Information Estimation of Simultaneous Equations Models with Lagged Endogenous Variables and Autoregressive Errors," Discussion Paper no. 183, The Wharton School of Finance and Commerce, University of Pennsylvania (Oct. 1970).

Fair, R. C., "The Estimation of Simultaneous Equation Models with Lagged Endogenous Variables and First Order Serially Correlated Errors," *Econometrica*, XXXVIII (May, 1970), 507–516.

Hendry, D. F., "Maximum Likelihood Estimation of Systems of Simultaneous Regression Equations with Errors Generated by a Vector Autoregressive Process," *International Economic Review*, XII (June, 1971).

Johnston, H. N., Ph.D. thesis, Department of Economics, University of Pennsylvania.

Sargan, J. D., "The Maximum Likelihood Estimation of Economic Relationships with Autoregressive Residuals," *Econometrica*, XXIX (July, 1961), 414–426.